Structure of the Wechsler Intelligence Scale for Children—Fourth Edition Among a National Sample of Referred Students

Marley W. Watkins
Arizona State University

The structure of the Wechsler Intelligence Scale for Children—Fourth Edition (WISC–IV; D. Wechsler, 2003a) was analyzed via confirmatory factor analysis among a national sample of 355 students referred for psychoeducational evaluation by 93 school psychologists from 35 states. The structure of the WISC–IV core battery was best represented by four first-order factors as per D. Wechsler (2003b), plus a general intelligence factor in a direct hierarchical model. The general factor was the predominate source of variation among WISC–IV subtests, accounting for 48% of the total variance and 75% of the common variance. The largest 1st-order factor, Processing Speed, only accounted for 6.1% total and 9.5% common variance. Given these explanatory contributions, recommendations favoring interpretation of the 1st-order factor scores over the general intelligence score appear to be misguided.

Keywords: intelligence, factor analysis, WISC–IV, children

Following substantial changes in content and structure (Kaufman, Flanagan, Alfonso, & Mascolo, 2006), the Wechsler Intelligence Scale for Children—Fourth Edition (WISC–IV; Wechsler, 2003a) replaced the Wechsler Intelligence Scale for Children—Third Edition (WISC–III; Wechsler, 1991) in 2003. For example, four of the 12 WISC–III subtests were omitted from the WISC–IV or made optional (Information, Picture Arrangement, Picture Completion, and Object Assembly), whereas three new subtests were added to the core WISC–IV (Picture Concepts, Matrix Reasoning, and Letter-Number Sequencing). In addition, artwork was revised and administration and scoring criteria were modified (Wechsler, 2003b). In total, approximately 60% of the items in the core WISC–IV subtests are new or revised.

Because test revisions "may assess traits, abilities, and conditions in ways different from earlier versions" (Strauss, Spreen, & Hunter, 2000, p. 237), psychometric standards (see American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) applicable to new tests should also govern test revisions (Adams, 2000). Of particular importance is evidence about the internal structure of the test revision, because there should be correspondence between the internal structure of the test and the structure of the construct assumed to be measured by the test (Messick, 1995).

Wechsler (2003b) reported that the WISC–IV was designed to measure intellectual functioning in four specific cognitive domains and to provide an overall composite that represents general ability. The WISC–IV structure was investigated with exploratory (EFA) and confirmatory (CFA) factor analyses of the normative sample with both core (10 subtest) and supplemental (15 subtest) batteries

(Wechsler, 2003b). Four first-order oblique factors were consistently identified: Verbal Comprehension (VC), Perceptual Reasoning (PR), Working Memory (WM), and Processing Speed (PS). These four dimensions were adopted as the scoring structure of the WISC–IV. The same structure was found in an EFA reanalysis of the normative data (Sattler, 2008), and measurement invariance across gender (H. Chen & Zhu, 2008) and age (Keith, Fine, Taub, Reynolds, & Kranzler, 2006) was subsequently demonstrated.

Inexplicably, Wechsler (2003b) did not conduct a higher order factor analysis to verify and describe the proposed multilevel structure of the WISC–IV. This oversight was addressed by three independent analyses of the WISC–IV standardization sample. In the first, Keith (2005) applied CFA to all 15 subtests but only considered two models that included a general intelligence factor. Both models exhibited good fit to the data. The second analysis applied CFA to all 15 WISC–IV subtests and concluded that a five-factor model based on the Cattell-Horn-Carroll (CHC; McGrew, 2009) theory might be the most appropriate model (Keith et al., 2006). However, Keith et al. (2006) did not analyze the WISC–IV core battery; the superiority of the CHC model over the Wechsler (2003b) four-factor model for the 15 subtest battery was not substantial in terms of fit (F. F. Chen, 2007); simple structure was abandoned in the CHC model; and the loading of the second-order *Gf* factor on the third-order general factor was 1.00, which was a replication of Gustafsson's (1984) results that questions the independence of these two factors. The third analysis applied the Schmid and Leiman (1957) orthogonal transformation of EFA results from the WISC–IV core battery to reveal the importance of the general intelligence factor in comparison to the four first-order factors (Watkins, 2006). Specifically, the general factor accounted for 71% of the common variance, whereas the largest first-order factor only contributed 12% of the common variance.

On the basis of these analyses, the factor structure of the WISC–IV among a nationally representative sample of children appears to consist of four to five first-order factors and a single

broad or superordinate factor associated with general intelligence. However, as one of the main uses of intelligence tests is the assessment of children with suspected disabilities, it must be demonstrated that the WISC–IV is appropriate for those children. Only two published studies have been conducted with a referral or clinical sample. In one study, a four-factor oblique EFA solution similar to that reported by Wechsler (2003b) was supported for the WISC–IV core battery in a sample of Pennsylvania students referred for evaluation to determine eligibility for special education services (Watkins, Wilson, Kotz, Carbone, & Babula, 2006). Similar to the results of Watkins (2006), transformation to an orthogonalized higher order model demonstrated that the general factor accounted for more than 75% of the common variance, whereas the largest first-order factor contributed only 10.5% of the common variance. The second study included 344 children with a variety of disorders who received a comprehensive neuropsychological evaluation, including the core WISC–IV battery, at a pediatric hospital (Bodin, Pardini, Burns, & Stevens, 2009). Similar to Keith (2005), a higher order four-factor model best fit the data of this clinical sample. The higher order general factor accounted for 77% of the common variance and, it is interesting to note, accounted for 99% of variance in the WM factor.

Although they are informative, these clinical studies were restricted to geographically local samples, and analytic methods were circumscribed. Additional studies with different clinical samples and varied methods are needed (Bodin et al., 2009; Strauss, Sherman, & Spreen, 2006), because the validity of a structural model is enhanced if the same model can be replicated in new samples (Raykov & Marcoulides, 2000). In addition, the structure of the core WISC–IV battery should receive further attention, because most clinicians only administer the 10 required subtests. For example, Watkins et al. (2006) found that fewer than 6% of their clinical sample received supplemental subtests. Consequently, the structure of the WISC–IV core battery was analyzed with CFA methods among a national sample of students referred for psychoeducational evaluation.

Method

Participants

Participants included 355 students (218 male and 137 female) who ranged in age from 6 to 16 years ($M = 9.78$ years, $SD = 2.54$). Grade placement ranged from kindergarten to Grade 11, with a median of fourth grade. Ethnic background of participants, if reported, was 62.7% White, 20.6% Black, 12.3% Hispanic, 2.4% Asian/Pacific Islander, and 2.0% other. All students were assessed to determine eligibility for special education services, but approximately 30% of the participants were determined not to be disabled. In contrast, 70% of the participants were reported to be eligible for special education services (41% with learning disabilities, 7% with mental retardation, 6% with emotional disabilities, 2% as gifted, 4% with speech disabilities, 9% with other health impairments, and 1% with autism spectrum disabilities). To ensure anonymity, no other demographic data were collected.

Instrument

The WISC–IV is a revised version of the WISC–III that was standardized on a nationally representative sample of 2,200 children aged 6–16 years closely approximating the 2000 U.S. Census on sex, race, parent education level, and geographic region. The WISC–IV core battery contains 10 core subtests ($M = 10$, $SD = 3$) that form the four (VC, PR, WM, and PS) factor indices ($M = 100$, $SD = 15$). The Full Scale IQ ($M = 100$, $SD = 15$) is based on the sum of scores from the 10 core subtests. Reliability and validity evidence was provided by Wechsler (2003b) and by Williams, Weiss, and Rolfhus (2003). Additional information on the WISC–IV was presented by Flanagan and Kaufman (2009) and by O'Donnell (2009).

Procedure

Following IRB approval, data for this study were solicited and collected electronically. A commercial marketing firm contacted 2,384 school psychologists via e-mail and invited them to anonymously contribute WISC–IV scores from anonymous students recently evaluated in their schools. Ninety-three (19 male and 74 female) school psychologists from 34 states responded by entering WISC–IV data for 355 students onto a Web form. On average, each school psychologist contributed 2.9 cases ($SD = 3.7$). Of the responding school psychologists, 19% held a master's degree, 65% a specialist degree, and 16% a doctoral degree. Years of experience of contributing psychologists ranged from 1 to 34 ($Mdn = 5.5$, $SD = 10.1$).

Analyses

CFA using maximum likelihood estimation was applied to the covariance matrix using EQS for Windows (Version 6.1). The obtained solutions were checked for convergence, and the adequacy of the parameter estimates and their associated standard errors were examined prior to considering the reported fit indices. According to Hu and Bentler (1998), values $\geq .95$ for comparative fit index (CFI), $\leq .08$ for standardized root-mean-square residual (SRMR), and $\leq .06$ for root-mean-square error of approximation (RMSEA) indicate that there is a good fit between the hypothesized model and the sample data. Although Marsh, Hau, and Wen (2004) cautioned against overgeneralization, these cutoff values seemed appropriate given the variables analyzed in the current study (Bodin et al., 2009). In addition, Akaike's information criterion (AIC) was consulted. The AIC considers statistical goodness-of-fit as well as model parsimony, with smaller values representing a better fit.

Each of the six models selected for this study was designed to evaluate a specific hypothesis about the structure of the WISC–IV (Bodin et al., 2009; Keith, 2005; Keith et al., 2006; Wechsler, 2003b). As illustrated in Table 2, the first four models contained one to four first-order oblique factors, as per Wechsler (2003b). The final two models conceptualized general intelligence in disparate ways, as per Keith (2005). The traditional higher order model specified that general intelligence had a direct influence on the first-order factors but only influenced the subtests indirectly through the first-order factors (see Figure 1). That is, the association between the general intelligence factor and the subtests was mediated by the first-order factors (Yung, Thissen, & McLeod, 1999). Gignac (2008) called this an indirect hierarchical model. In contrast, the direct hierarchical model (see Figure 2) specified a first-order general intelligence factor that had a direct effect on the
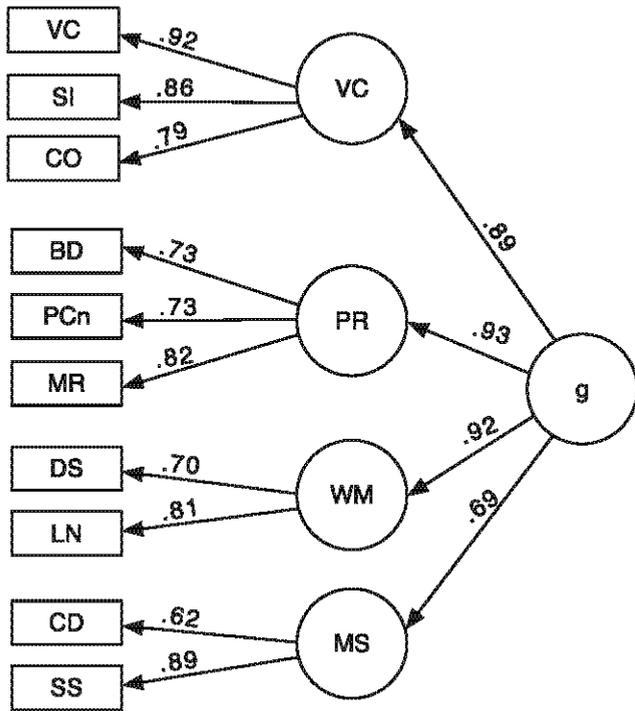
Figure 1. Indirect hierarchical structure, with standardized coefficients, of the Wechsler Intelligence Scale for Children—Fourth Edition (Wechsler, 2003a) for 355 referral students. SI = Similarities; VC = Vocabulary; CO = Comprehension; BD = Block Design; PCn = Picture Concepts; MR = Matrix Reasoning; DS = Digit Span; LN = Letter-Number Sequencing; CD = Coding; SS = Symbol Search; VC = Verbal Comprehension; PR = Perceptual Reasoning; WM = Working Memory; PS = Processing Speed; g = General Intelligence.

10 subtests. Each of the four first-order factors also had a direct effect on specific subtests, but the general factor had no effect on the first-order factors. Thus, each subtest contributed variance directly to the general intelligence factor and to a narrower group factor. This model has variously been called the bi-factor model (Holzinger & Swineford, 1937), nested factor model (Gustafsson & Balke, 1993), and direct hierarchical model (Gignac, 2008). These final two models contrasted general intelligence as a higher order superordinate factor versus a first-order breadth factor, respectively (Gignac, 2008).

Results

Participants' mean WISC–IV subtest, factor, and IQ scores were slightly lower and somewhat more variable than the normative sample (see Table 1). Similar patterns have been found with other samples of referred students (Canivez & Watkins, 1998). Nevertheless, score distributions appeared to be relatively normal, with .51 the largest skew and .89 the largest kurtosis (Onwuegbuzie & Daniel, 2002). Two other conditions for multivariate normality are that all linear combinations of variables follow a normal distribution, and all subsets of variables in the data set are normally distributed (Stevens, 2009). This was verified by examining the scatterplots of all variable pairs. All scatterplots had an elliptical shape. In addition, multivariate kurtosis was examined with the
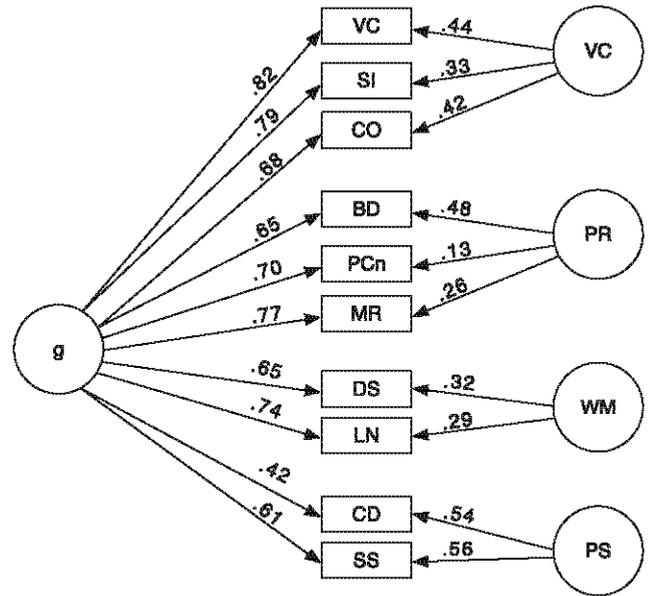

Figure 2. Direct hierarchical structure, with standardized coefficients, of the Wechsler Intelligence Scale for Children—Fourth Edition (Wechsler, 2003a) for 355 referral students. SI = Similarities; VC = Vocabulary; CO = Comprehension; BD = Block Design; PCn = Picture Concepts; MR = Matrix Reasoning; DS = Digit Span; LN = Letter-Number Sequencing; CD = Coding; SS = Symbol Search; VC = Verbal Comprehension; PR = Perceptual Reasoning; WM = Working Memory; PS = Processing Speed; g = General Intelligence.

normalized value calculated by EQS. In practice, normalized values greater than |5.0| indicate that the data are multivariate nonnormal (Byrne, 2006). The normalized value for this sample was 3.78, indicating a reasonable degree of multivariate normality.

A review of model fit statistics (see Table 3) indicates that the final three models met a priori guidelines for a close fit (i.e., CFI ≥ .95, SRMR ≤ .08, and RMSEA ≤ .06). Thus, the one-, two-, and

Table 1

Descriptive Statistics for 355 Referral Students on the Wechsler Intelligence Scale for Children—Fourth Edition

WISC–IV component	M	SD	Skew	Kurtosis
Verbal Comprehension Index	92.4	16.5	+.30	+.86
Perceptual Reasoning Index	95.5	16.9	−.04	+.02
Working Memory Index	89.7	15.6	+.11	+.52
Processing Speed Index	90.8	15.3	+.06	+.08
Full Scale IQ	90.6	16.9	+.09	+.76
Block Design	8.9	3.3	+.23	−.36
Similarities	8.8	3.4	+.35	+.03
Digit Span	8.1	3.0	+.24	+.44
Picture Concepts	9.7	3.3	−.26	−.03
Coding	8.3	3.2	+.35	+.09
Vocabulary	8.6	3.1	+.27	+.53
Letter-Number Sequencing	8.4	3.3	−.33	+.02
Matrix Reasoning	9.0	3.3	+.28	+.04
Comprehension	8.7	3.2	−.05	−.02
Symbol Search	8.3	3.1	−.33	+.23

Note. WISC–IV = Wechsler Intelligence Scale for Children—Fourth Edition (Wechsler, 2003a).

Table 2
Six Structural Models of the Wechsler Intelligence Scale for Children—Fourth Edition

Model	Subtest									
	SI	VC	CO	BD	PCn	MR	DS	LN	CD	SS
One factor	I	I	I	I	I	I	I	I	I	I
Two factor	I	I	I	II	II	II	I	I	II	II
Three factor	I	I	I	II	II	II	III	III	III	III
Four factor	I	I	I	II	II	II	III	III	IV	IV
Indirect hierarchical	I	I	I	II	II	II	III	III	IV	IV
Direct hierarchical	I	I	I	II	II	II	III	III	IV	IV

Note. SI = Similarities; VC = Vocabulary; CO = Comprehension; BD = Block Design; PCn = Picture Concepts; MR = Matrix Reasoning; DS = Digit Span; LN = Letter-Number Sequencing; CD = Coding; SS = Symbol Search; I is factor one, II is factor II, III is factor III, and IV is factor four. Wechsler Intelligence Scale for Children—Fourth Edition (Wechsler, 2003a).

three-factor models were inadequate. RMSEA and AIC values for the first-order four-factor model were inferior to those of the hierarchical models. Beyond statistical fit, the oblique four-factor model was unsatisfactory because it did not include general intelligence as specified by Wechsler (2003b) and others (Carroll, 1993, 2003; Gustafsson, 1994; Jensen & Weng, 1994). Of the two remaining models, the direct hierarchical model was statistically superior to the indirect hierarchical model ($df = 2$, $\Delta\chi^2 = 6.68$, $p = .048$).

On the basis of these statistical and theoretical arguments, the direct hierarchical model was deemed preferable and is presented with standardized loadings in Figure 2. Overall, the influence of general intelligence dwarfed the contributions made by the four WISC–IV first-order factors. The general factor accounted for 75% of the common variance and 48% of the total variance. The VC factor accounted for 7.5% common and 4.8% total variance, the PR factor for 4.9% common and 3.1% total variance, the WM factor for 2.9% common and 1.9% total variance, and the PS factor for 9.5% common and 6.1% total variance. Altogether, the general and broad factors accounted for approximately 64% of the total variance, leaving 36% unique variance.

Discussion

Factor analyses of the WISC–IV scores of a national sample of students referred to school psychologists for evaluation to determine eligibility for special education services essentially replicated the results of Wechsler (2003b), Keith (2005), and Watkins (2006)

for the general population as well as Watkins et al. (2006) and Bodin et al. (2009) for clinical samples. The structure of the WISC–IV core battery among this sample of students was best represented by the four first-order factors named VC, PR, WM, and PS by Wechsler (2003b), plus a general intelligence factor.

Analyses of the Wechsler Adult Intelligence Scale—Revised (Wechsler, 1981) and the Wechsler Adult Intelligence Scale—Third Edition (Wechsler, 1997) favored a direct hierarchical model (Gignac, 2005, 2006). The direct hierarchical model was also a statistically better fit than an indirect hierarchical model with the WISC–IV normative sample, but Keith (2005) preferred the indirect hierarchical model on theoretical grounds. However, it is difficult to interpret the relative contributions of first-order and higher order factors within the indirect hierarchical model without additional computations or transformations (e.g., a Schmid and Leiman, 1957, orthogonalization as recommended by Carroll, 1993). In addition, Gignac (2006) questioned the appropriateness of the full mediation implied by the indirect hierarchical model and suggested that "it is arguably more congruent and reasonable to specifically model the most significant factor of a battery of tests (i.e., 'g') directly, rather than indirectly, through first-order factors" (p. 85). Gignac (2008) also reasoned that a first-order breadth factor is more appropriate than a higher order superordinate factor, because a principal aspect of general intelligence is its breadth rather than its superordination. Following this logic, Gustafsson (2001) observed that the direct hierarchical model often "results in more parsimonious models with fewer latent variables" (p. 30).

Table 3
Fit Statistics for Six Structural Models of the Wechsler Intelligence Scale for Children—Fourth Edition Among 355 Students Referred for Psychoeducational Evaluation

Model	df	χ^2	CFI	SRMR	RMSEA	90% CI RMSEA	AIC
One factor	35	240.87	.893	.063	.129	.114–.144	170.87
Two factor	34	182.80	.922	.059	.111	.095–.127	114.80
Three factor	32	120.02	.954	.046	.088	.071–.105	56.02
Four factor	29	65.59	.981	.028	.060	.040–.079	7.59
Indirect hierarchical	31	68.48	.980	.030	.058	.040–.077	6.48
Direct hierarchical	27	58.91	.983	.028	.058	.038–.078	4.91

Note. CFI = comparative fit index; SRMR = standardized root-mean-square residual; RMSEA = root-mean-square error of approximation; CI = confidence interval; AIC = Akaike's information criterion. All χ^2 values $p < .001$. Loadings of two-subtest factors constrained to be equal to ensure local independence in the direct hierarchical model. Wechsler Intelligence Scale for Children—Fourth Edition (Wechsler, 2003a).

F. F. Chen, West, and Sousa (2006) also noted that the direct hierarchical model offers greater utility in predicting external criteria. This benefit was theoretically and empirically explicated by Schmiedek and Li (2004) and Brunner (2008). A linear dependency of parameters imposed by the structure of the indirect hierarchical model "does not afford simultaneous estimations of general and specific effects" (Schmiedek & Li, 2004), and results in "a severe limitation to the understanding of how general and specific abilities relate to other theoretical constructs" (p. 162). Thus, the WISC–IV general intelligence factor is best interpreted as a first-order breadth factor as specified in the direct hierarchical model (see Figure 2).

The general factor was the predominate source of variation among WISC–IV subtests, accounting for 75% of the common variance and 48% of the total variance. In fact, the general factor explained three times the variance of the four narrow factors combined. As a comparison, the general factor accounted for 75.7% of the common variance among a sample of referred students in Pennsylvania (Watkins et al., 2006), 77.2% of the common variance among a pediatric neuropsychological sample in the southeastern United States, and 71.3% of the common variance among the WISC–IV normative sample (Watkins, 2006). Given these explanatory contributions, recommendations favoring interpretation of the narrow factor scores over the general intelligence score (Prifitera, Saklofske, & Weiss, 2008; Wechsler, 2003b; Weiss, Saklofske, Prifitera, & Holdnack, 2006; Williams et al., 2003) appear to be misguided.

As with all studies, these conclusions must be tempered by limitations of sampling and data collection. Although 93 school psychologists contributed WISC–IV data, this was only 4% of those solicited for participation. Other studies have also found low response rates to such requests (Canivez & Watkins, 1998), and there is no way to know if this particular sample was representative of the population. Likewise, the use of electronic solicitation and data collection may have affected the data in unknown ways. Fortunately, these concerns are mitigated by results that were consistent with previous studies of the nationally representative standardization sample and two clinical samples. Therefore, the best available evidence indicates that clinicians should favor interpretation of the general intelligence score over the narrow factor scores when using the WISC–IV core battery.

References

Adams, K. M. (2000). Practical and ethical issues pertaining to test revisions. *Psychological Assessment, 12,* 281–286.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). S*tandards for educational and psychological testing.* Washington, DC: Author.

Bodin, D., Pardini, D. A., Burns, T. G., & Stevens, A. B. (2009). Higher order factor structure of the WISC–IV in a clinical neuropsychological sample. *Child Neuropsychology, 15,* 417–424.

Brunner, M. (2008). No *g* in education? *Learning and Individual Differences, 18,* 152–165.

Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Erlbaum.

Canivez, G. L., & Watkins, M. W. (1998). Long-term stability of the Wechsler Intelligence Scale for Children—Third Edition. *Psychological Assessment, 10,* 285–291.

Carroll, J. B. (1993). *Human cognitive abilities.* Cambridge, England: Cambridge University Press.

Carroll, J. B. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports *g* and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 5–21). New York, NY: Pergamon.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464–504.

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research, 41,* 189–225.

Chen, H., & Zhu, J. (2008). Factor invariance between genders of the Wechsler Intelligence Scale for Children—Fourth Edition. *Personality and Individual Differences, 45,* 260–266.

Flanagan, D. P., & Kaufman, A. S. (2009). *Essentials of WISC–IV assessment* (2nd ed.). Hoboken, NJ: Wiley.

Gignac, G. E. (2005). Revisiting the factor structure of the WAIS–R: Insights through nested factor modeling. *Assessment, 12,* 320–329.

Gignac, G. E. (2006). The WAIS–III as a nested factors model: A useful alternative to the more conventional oblique and higher-order models. *Journal of Individual Differences, 27,* 73–86.

Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: *g* as superordinate or breadth factor. *Psychology Science Quarterly, 50,* 21–43.

Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8,* 179–203.

Gustafsson, J.-E. (1994). Hierarchical models of intelligence and educational achievement. In A. Demetriou & A. Efklides (Eds.), *Intelligence, mind, and reasoning: Structure and development* (pp. 45–73). New York, NY: Elsevier.

Gustafsson, J.-E. (2001). On the hierarchical structure of ability and personality. In J. M. Collis & S. Messick (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement* (pp. 25–42). Mahwah, NJ: Erlbaum.

Gustafsson, J.-E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research, 28,* 407–434.

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika, 2,* 42–54.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424–453.

Jensen, A. R., & Weng, L.-J. (1994). What is good *g*? *Intelligence, 18,* 231–258.

Kaufman, A. S., Flanagan, D. P., Alfonso, V. C., & Mascolo, J. T. (2006). [Review of the Wechsler Intelligence Scale for Children, Fourth Edition (WISC–IV)]. *Journal of Psychoeducational Assessment, 24,* 278–295.

Keith, T. Z. (2005). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 581–614). New York, NY: Guilford Press.

Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order, multisample, confirmatory factor analysis of the Wechsler Intelligence Scale for Children—Fourth Edition: What does it measure? *School Psychology Review, 35,* 108–127.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11,* 320–341.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37,* 1–10.

Messick, S. (1995). Validity of psychological assessment. *American Psychologist, 50,* 741–749.

O'Donnell, L. (2009). The Wechsler Intelligence Scale for Children—Fourth Edition. In J. A. Naglieri & S. Goldstein (Eds.), *Practitioner's guide to assessing intelligence and achievement* (pp. 153–190). Hoboken, NJ: Wiley.

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools, 9,* 73–90.

Prifitera, A., Saklofske, D. H., & Weiss, L. G. (2008). *WISC–IV clinical use and interpretation: Scientist-practitioner perspectives* (2nd ed.). San Diego, CA: Academic Press.

Raykov, T., & Marcoulides, G. A. (2000). *A first course in structural equation modeling.* Mahwah, NJ: Erlbaum.

Sattler, J. M. (2008). *Assessment of children: Cognitive foundations* (5th ed.). San Diego, CA: Author.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22,* 53–61.

Schmiedek, F., & Li, S.-C. (2004). Toward an alternative representation for disentangling age-associated differences in general and specific cognitive abilities. *Psychology and Aging, 19,* 40–56.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York, NY: Routledge.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary.* New York, NY: Oxford University Press.

Strauss, E., Spreen, O., & Hunter, M. (2000). Implications of test revisions for research. *Psychological Assessment, 12,* 237–244.

Watkins, M. W. (2006). Orthogonal higher order structure of the Wechsler Intelligence Scale for Children—Fourth Edition. *Psychological Assessment, 18,* 123–125.

Watkins, M. W., Wilson, S. M., Kotz, K. M., Carbone, M. C., & Babula, T. (2006). Factor structure of the Wechsler Intelligence Scale for Children—Fourth Edition among referred students. *Educational and Psychological Measurement, 66,* 975–983.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised.* New York, NY: Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—Third Edition.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003a). *Wechsler Intelligence Scale for Children—Fourth Edition.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003b). *Wechsler Intelligence Scale for Children—Fourth Edition technical and interpretive manual.* San Antonio, TX: Psychological Corporation.

Weiss, L. G., Saklofske, D. H., Prifitera, A., & Holdnack, J. A. (2006). *WISC–IV advanced clinical interpretation.* San Diego, CA: Academic Press.

Williams, P. E., Weiss, L. G., & Rolfhus, E. L. (2003). *Clinical validity* (WISC–IV Technical Report #3). San Antonio, TX: Psychological Corporation.

Yung, Y.-F., Thissen, D., & McLeod, L. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika, 64,* 113–128.